XP-002165267

TECHNICAL ADVANCE

Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs

Michèle Minet*, Marie-Elisabeth Dufour and François Lacroute
Centre de Génétique Moléculaire, CNRS F 91190, Gif sur Yvette, France

Summary

An *Arabidopsis thaliana* cDNA bank has been constituted in a *Saccharomyces cerevisiae* expression vector based on the phosphoglycerate kinase (PGK) promoter and terminator. This bank was used to complement eight *S. cerevisiae* auxotrophic markers. All of them were corrected. These results confirm the quality of the bank and the feasibility of cloning plant genes by yeast mutant complementation. The cDNA complementing the *ura1* yeast mutant was sequenced, analysed and shown to encode a dihydroorotic (DHO) dehydrogenase sequence.

Introduction

Since the cloning of a human cDNA by direct complementation of the fission yeast *Schizosaccharomyces pombe* (Lee and Nurse, 1987), many other genes have been obtained by this approach (see Becker *et al.*, 1991; Lew *et al.*, 1991 and references therein; Minet and Lacroute, 1990; Schild *et al.*, 1990). There are two main limitations to this technique. First, only the basic functions of a eukaryotic cell are expressed in yeast, which excludes cloning the genes involved in differentiation. Secondly, many proteins function in narrow interaction with neighbouring proteins or nucleic acids (RNA polymerase subunits, ribosomal proteins, proteins involved in RNA splicing, etc.). In this case, heterospecific complementation is unlikely to function effectively in organisms which have widely diverged in evolution. We have prepared an *A. thaliana* cDNA bank in a yeast expression vector to test if the yeast complementation cloning approach, which was successful with mammalian genes, would work equally well with plants.

Results

Construction of the yeast vector

We previously constructed pFL60 (Minet and Lacroute, 1990), a yeast cDNA cloning vector adapted to the direct

cloning of cDNAs on the oligo(dT) tail borne by the vector itself (see Okayama and Berg, 1982). This vector contains the phosphoglycerate kinase (PGK) promoter separated from its terminator by a multisite polylinker, the yeast *URA3* gene, a small part of the 2 µm yeast plasmid bearing the replication origin and the *E. coli* vector pUC19. An oligo(dT) adaptor subsequently had to be introduced in a unidirectional manner. As the Okayama and Berg method gave us too few clones, we decided to adapt pFL60 for the use of the *Bst*XI double-site digestion technique which creates incompatible cohesive ends on the vector (Aruffo and Seed, 1987). A *Not*I linker was introduced in place of the *Cla*I–*Bgl*II polylinker of pFL60 and a new polylinker was subsequently introduced in the *Not*I site with the sequence given in Figure 1. This new vector was named pFL61.

Preparation of the cDNA bank

Complete young *A. thaliana* seedlings (stage two leaves), including roots, were frozen in liquid nitrogen and ground
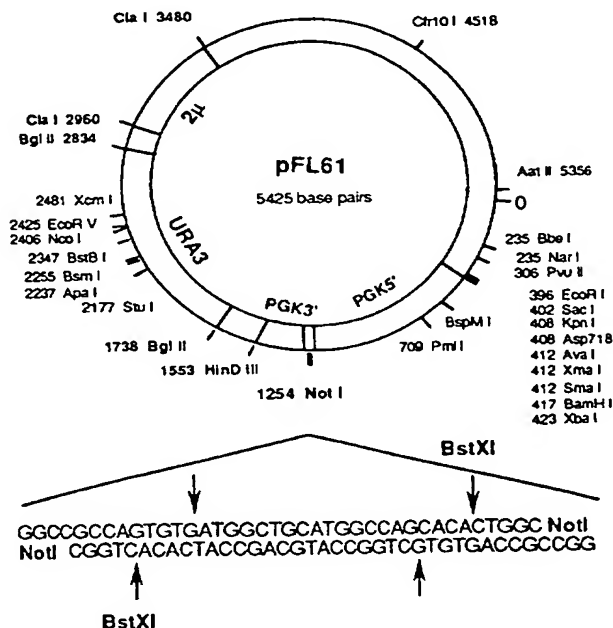

Figure 1. Restriction map of the pFL61 yeast expression vector and sequence of the polylinker introduced in the *Not*I site.
The restriction sites indicated are one cut sites; the *Cla*I and *Bgl*II sites were used for the construction of pFL61. *Bst*XI sites are the insertion sites for the cDNA pool.

417

in a mortar with frozen TE buffer with 5% SDS. The resulting powder was melted, extracted twice with an equal volume of phenol–chloroform and twice again after addition of 1/10 volume of 4 M NaCl. Nucleic acids were precipitated with two volumes of ethanol and immediately enriched in poly(A)$^+$ RNAs by chromatography on an oligo(dT) column. After elution and ethanol precipitation, the poly(A)$^+$ RNAs were resuspended in water, quantified by UV absorbance and immediately frozen at −80°C until use. The cDNAs were prepared from 5 μg of poly(A)$^+$ RNAs with an Amersham 'cDNA plus' cloning kit, using AMV reverse transcriptase for the synthesis of the first strand and RNase H and the Klenow fragment of *E. coli* DNA polymerase for the second strand. To quantify subsequent cDNA recovery, a small amount (1 μCi) of radioactive dCTP was added during the synthesis of the second strand. After blunt-end conversion, the cDNAs were ligated to DNA adaptors with CACA cohesive ends compatible with the GTGT ends produced in the vector by *Bst*XI digestion. The cDNAs were then separated by size on a low melting point agarose gel which also removed the excess DNA adaptors. Three size classes were recovered from the gel: one from 0.2 to 1 kb, one from 1 to 2.5 kb and the last one from 2.5 kb to the top of the gel. The calculated recovery (as a percentage of the total radioactivity) of the cDNAs was 33% in the short class, 22% in the medium class and 10% in the large class; a rough evaluation of the corresponding amounts of cDNAs gave, respectively, 1.5 μg, 1 μg and 0.4 μg. An approximately stoichiometric amount of *Bst*XI digested vector was added (with prior removal of the small internal *Bst*XI fragment) and each class was ligated separately. After *E. coli* transformation by electroporation (strain MR 32), 7 × 10^5 clones were obtained from the short class, 10^6 clones from the medium class and 5.6 × 10^5 clones from the longer class. Eight clones, chosen randomly in each class size, were analysed and the length of their cDNA insert measured after digestion by the restriction enzyme *Not*I. The average insert size is 0.7 kb for the short class, 1.6 kb for the medium class and 2.4 kb for the long class. The clones were then pooled together, an aliquot was frozen at −80°C for future amplification and the remainder was used to prepare plasmid DNA in a caesium chloride gradient.

Complementation of *S. cerevisiae* mutants

To test the cDNA bank, different auxotrophic strains of *S. cerevisiae* were transformed with purified plasmid DNA and prototrophic clones were selected for growth on appropriate medium. The strains used, the transformation method and the selection details are given in Experimental procedures. Eight different genes were tested for complementation: *URA1, URA2, URA4, URA5–URA10, ADE2, HIS3, LEU2* and *TRP1*; all of the eight yeast mutants
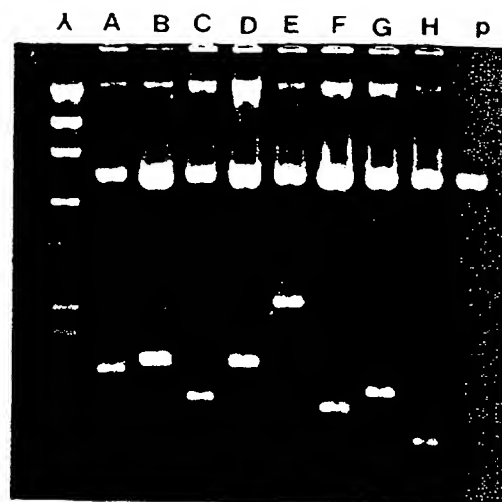

Figure 2. Electrophoresis after *Not*I digestion of *A. thaliana* cDNA inserts complementing: A, *ura1*; B, *ura2*; C, *ura4*; D, *ura5–ura10*; E, *ade2*; F, *his3*; G, *leu2*; H, *trp1 S. cerevisiae* mutants. λ: phage λ digested by *Hin*dIII as molecular weight standard. p: pFL61 digested by *Not*I.

were complemented by the pool. It has been shown (Bertauche *et al.* unpublished results) that the *ura5–ura10* complementing cDNA was also able to complement a *ura3* mutant indicating that, as expected from enzymological results (Doremus, 1986; Walther *et al.*, 1984), plants possess a bifunctional protein having both orotate phosphoribosyl transferase and orotidine 5′-phosphate decarboxylase activities.

In the case of *ade2, his3, leu2* and *trp1* complementation, the yeast strain also harboured a *ura3* marker, which allowed us to measure the total number of transformants (complementation of *ura3* by the *URA3* gene of the pFL61 plasmid) and thus estimate the frequency of complementing cDNA, which ranged between 5 × 10^{-5} and 5 × 10^{-4}. A similar number of complementing clones was observed in the case of *ura4* and of *ura5–ura10* (where the total number of transformants could not be measured). However this number was one and two orders of magnitude lower for the *ura1* and *ura2* complementation, respectively. Moreover, the *ura1* and *ura2* complementing clones grew very poorly (doubling time of 5 h on minimal medium compared to 2 h for a prototrophic wild-type strain), indication a poor complementation.

In each case, the occurrence of true complementation was confirmed by a loss of prototrophic phenotype accompanying loss of the plasmid, and by a systematic recovery of the prototrophic phenotype after retransformation by the purified plasmid.

Figure 2 presents the size of the different complementing inserts after *Not*I digestion. In view of our previous interest in the pyrimidine pathway we have chosen to study the *ura1* complementing insert in more detail. The corresponding plasmid was called p*PyrD ARA*.

```
                                                        -90
                                               CTAGAAGCTTA
AAACCCTCTTCTTATAGATTTCAAAGCAACCATTTTCGAACTCTGTTTGATTCCACTGTCGATTTAATCATCTTACCTA
1/1                                          31/11
ATG GCC GGA AGG GCT GCG ACG TCG TCG GCG AAA TGG GCG AGA GAG TTT TTG TTC AGA AGG
Met ala gly arg ala ala thr ser ser ala lys trp ala arg glu phe leu phe arg arg
61/21                                        91/31
GTT TCG TCT AAT CCT CTT GGA GCT ACC CGT AAC TGT TCT TCG GTT CCT GGA GCT TCT TCT
val ser ser asn pro leu gly ala thr arg asn cys ser ser val pro gly ala ser ser
121/41                                       151/51
GCA CCG AAA GTC CCT CAT TTT TCC AAG AGA GGA AGG ATA TTG ACA GGA GCT ACC ATT GGT
ala pro lys val pro his phe ser lys arg gly arg ile leu thr gly ala thr ile gly
181/61                                       211/71
CTG GCC ATA GCT GGA GGA GCT TAT GTT AGT ACT GCA GAT GAA GCA ACC TTC TGT GGG TGG
leu ala ile ala gly gly ala tyr val ser thr ala asp glu ala thr phe cys gly trp
241/81                                       271/91
CTA TTC AAC GCA ACA AAG GTT GTG AAT CCT TTC TTT GCT CTT CTG GAT GCT GAG TTT GCG
leu phe asn ala thr lys val val asn pro phe phe ala leu leu asp ala glu phe ala
301/101                                      331/111
CAT AAG CTG GCT GTT TCA GCT GCA GCT CGC GGG TGG GTA CCT AGA GAA AAG AGG CCT GAT
his lys leu ala val ser ala ala ala arg gly trp val pro arg glu lys arg pro asp
361/121                                      391/131
CCA GCA ATC TTG GGA CTT GAA GTT TGG GGA AGG AAA TTC TCA AAC CCA ATT GGG CTT GCT
pro ala ile leu gly leu glu val trp gly arg lys phe ser asn pro ile gly leu ala
421/141                                      451/151
GCT GGA TTT GAC AAA AAC GCT GAG GCT ACT GAG GGA CTG CTA GGG ATG GGA TTT GGC TTT
ala gly phe asp lys asn ala glu ala thr glu gly leu leu gly met gly phe gly phe
481/161                                      511/171
GTT GAG GTA GGC TCT GTA ACT CCT GTT CCA CAA GAA GGC AAT CCG AAA CCA CGT ATC TTC
val glu val gly ser val thr pro val pro gln glu gly asn pro lys pro arg ile phe
541/181                                      571/191
AGA CTA AGC CAA GAA GGA GCT ATT ATC AAT AGG TGT GGA TTT AAT AGC GAA GGG ATT GTT
arg leu ser gln glu gly ala ile ile asn arg cys gly phe asn ser glu gly ile val
601/201                                      631/211
GTC GTT GCA AAG CGG TTG GGT GCT CAA CAT GGT AAA AGA ATG TTG GCT GAA ACA TCA GCT
val val ala lys arg leu gly ala gln his gly lys arg met leu ala glu thr ser ala
661/221                                      691/231
ACT TCG TCA TCT CCA AGT GAT GAT GTA AAA CCA GGG GGA AAA TCT GGA CCT GGT ATT CTT
thr ser ser ser pro ser asp asp val lys pro gly gly lys ser gly pro gly ile leu
721/241                                      751/251
GGG GTC AAC CTT GGA AAG AAC AAG ACG AGT GAG GAT GCT GCT GCT GAC TAT GTC CAA GGA
gly val asn leu gly lys asn lys thr ser glu asp ala ala ala asp tyr val gln gly
781/261                                      811/271
GTT CAT AAC TTA TCC CAG TAT GCT GAT TAC TTG GTG ATC AAT GTT TCA TCA CCC AAT ACT
val his asn leu ser gln tyr ala asp tyr leu val ile asn val ser ser pro asn thr
841/281                                      871/291
GCA GGA CTG CGC ATG CTT CAA GGA AGG AAA CAG TTG AAG GAC CTT GTA AAG AAG GTT CAA
ala gly leu arg met leu gln gly arg lys gln leu lys asp leu val lys lys val gln
901/301                                      931/311
GCT GCT CGG GAT GAG ATG CAA TGG GGT GAT GAA GGT CCT CCT CCT CTT CTT GTG AAG ATT
ala ala arg asp glu met gln trp gly asp glu gly pro pro pro leu leu val lys ile
961/321                                      991/331
GCT CCT GAT CTG TCC AGA GGA GAG CTT GAA GAT ATT GCA GCG GTG GCT CTT GCT CTG CAC
ala pro asp leu ser arg gly glu leu glu asp ile ala ala val ala leu ala leu his
1021/341                                     1051/351
TTA GAT GGG CTG ATC ATA TCA AAT ACA ACA GTC TCA AGG CCT GAT GCT GTA AGC AAC AAC
leu asp gly leu ile ile ser asn thr thr val ser arg pro asp ala val ser asn asn
1081/361                                     1111/371
CCC GTC GCA ACA GAA ACA GGT GGT TTG AGC GGG AAA CCG CTC TTT GCT CTC TCC ACC AAC
pro val ala thr glu thr gly gly leu ser gly lys pro leu phe ala leu ser thr asn
1141/381                                     1171/391
ATG TTG AGA GAT ATG TAC ACT TTG ACA CGA GGA AAG ATT CCA TTG ATA GGC TGT GGC GGG
met leu arg asp met tyr thr leu thr arg gly lys ile pro leu ile gly cys gly gly
1201/401                                     1231/411
GTT AGT AGT GGT GAG GAT GCT TAC AAG AAA ATA AGA GCT GGA GCT ACT CTT GTT CAG CTG
val ser ser gly glu asp ala tyr lys lys ile arg ala gly ala thr leu val gln leu
1261/421                                     1291/431
TAC ACG GGA TTT GCC TAT GGA GGA CCT GCC CCA TCC CAC AAA TAA AGGAGGAACTGGTGAAATG
tyr thr gly phe ala tyr gly gly pro ala pro ser his lys OCH

CTTAGAAAGGGATGGCTTCAAGTCGATCCATGAAGCAATTGGTGCTGATCACAGATGATAAAACTCAAAGAGCAATACG
CTGCGGAAAGATCCCTTTTCTTAAAACAACAAAACAAGTTGTATTTAAAGTTCATTGATTCAGTTGTCTCTTTCTACCC
CCAACAATAACATCATCGTTTTTGCC
                    1508
```

Figure 3. Nucleotide and deduced amino acid sequences of the 1.6 kb cDNA insert. The possible polyadenylation signal AATAAC is underlined.

Sequence analysis

The 1598 bp sequence of the *PyrD ARA* insert is given in Figure 3 (Genbank/EMBL accession number X62909). There is a single open reading frame from base 91 to base 1395, corresponding to a 435 amino acid coding sequence from the first ATG onward and followed by a 203 bp tail with a potential adenylation signal (AAUAAC) (Wickens and Stefenson, 1984) 16 bp before the 3′ end. No poly(A) tail was found, which is probably due to a reduction in the length of the poly(dA/dT) tail during replications in *E. coli*.

Comparison with the NBRF protein databank reveals homology with the product of four genes known to encode DHO dehydrogenase (EC 1.3.3.1) in *Salmonella typhimurium* (Frick *et al.*, 1990), *E. coli* (Larsen and Jensen, 1985), *S. cerevisiae* and *Dictyostelium discoideum* (Jacquet *et al.*, 1985). These sequences were compared by dot matrix analysis (Figure 4) and by local alignment with the LFASTA program (data not shown). These comparisons revealed that the first 130 amino acids are specific to *A. thaliana* and that a small central sequence between 210 and 240 amino acids is present in *A. thaliana* and absent in the bacterial sequences. Statistical analysis of the simi-

larities between pairs of sequences using a ktup value of 1 and 100 shuffled sequences, gave significant *Z* values of 34.7, 21.4, 8.6 and 5.7, respectively (see Lipman and Pearson, 1985 and Pearson and Lipman, 1988 for discussion of the statistical meaning of the *Z* value).

The 58 amino acid N-terminal sequence has features typical of mitochondrial targeting signals, i.e. it is rich in positively charged and hydroxylated residues and lacks acidic residues (Hartl *et al.*, 1989). This is consistent with biochemical data (Doremus and Jagendorf, 1985) indicating that plant DHO dehydrogenase is a mitochondrial enzyme.

Northern analysis

Poly(A$^+$) RNAs from *A. thaliana* seedlings were analysed by Northern blotting, using the *PyrD* cDNA insert obtained after *Not*I digestion as a DNA probe. Two *A. thaliana* mRNAs of about 1.8 kb and 3.4 kb were detected. The 3.4 kb mRNA was about 10 times less abundant than the 1.8 kb species. No signal was obtained in the yeast RNA preparation (strain FL100) used as a control (Figure 5).

Discussion

We have constructed an *A. thaliana* cDNA bank in the yeast expression vector pFL61. This bank is very efficient in cloning plant cDNAs by complementation of yeast mutants. All the eight yeast mutants tested could be complemented by this bank. This success may be due to the fact that we restricted our assays to enzymes involved in biosynthetic pathways. However, cDNAs encoding the p450 reductase proteins (Pompon, personal communication) and the K$^+$ transport protein (Sentenac *et al.*, 1992) have also been obtained from this bank by yeast mutant
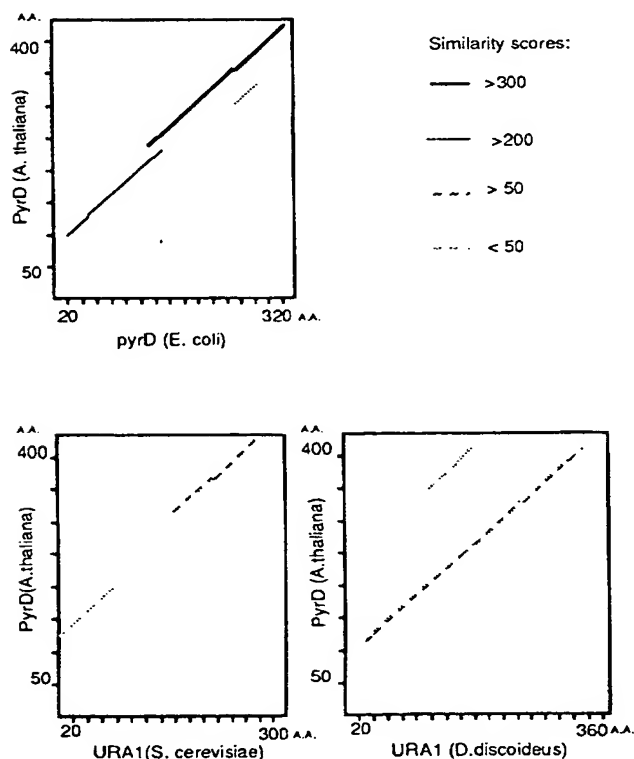

Figure 4. Dot matrix comparisons of DHO dehydrogenase amino acid sequences.
Matrix of Dayhoff, ktup = 1, 100 shuffled sequences. Sources: *E. coli* (Larsen and Jensen, 1986). Similar results were obtained with *S. typhimurium* (Frick *et al.*, 1990), *D. discoideum* (Jacquet *et al.*, 1985) and *S. cerevisiae* (EMBL accession number F32729).
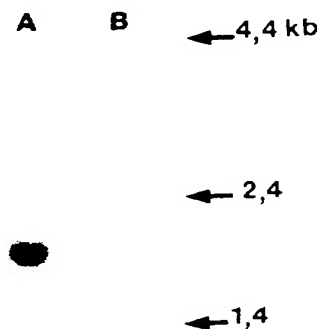

Figure 5. Northern analysis of poly(A)⁺ RNAs.
Lane A. *A. thaliana*, lane B, *S. cerevisiae*.

complementation. Furthermore, Alderson *et al.* (1991) have recently complemented an *sfn1* mutant by a rye cDNA coding for a protein kinase.

The poor complementation of the *ura2* mutant lacking aspartate transcarbamylase may reflect a difference in the synthesis of carbamyl phosphate (CP) and its regulation, in agreement with previous work indicating that in plants there is only one CP synthetase common to the arginine and pyrimidine pathways (O'Neal and Naylor, 1976). The 1.6 kb cDNA insert (*PyrD ARA*) which complements the *ura1* yeast mutant was analysed by Northern blot and sequencing. This cDNA hybridizes with two mRNAs of about 1.8 and 3.4 kb. The cDNA itself is only 1.6 kb long rather than 1.8 kb, which probably results from the shortening of the poly(dA/dT) tail during replication in *E. coli*. The 3.4 kb mRNA has a low abundance. The simplest explanation is that both transcripts are synthesized from the same gene by alternative splicing. The existence of an homologous gene appears unlikely since only one *Hin*dIII DNA fragment (2.4 kb) is detected at high stringency, in a Southern blot experiment (data not shown).

Amino acid sequence comparisons have shown that the plant DHO dehydrogenase displays much more sequence similarity to the bacterial enzyme than to the DHO dehydrogenase of eukaryotes such as *S. cerevisiae* and *D. discoideum* (Jacquet *et al.*, 1985). One explanation would be the occurrence of intergenic gene transfer between plants and bacteria. However, a more realistic hypothesis is convergent evolution where the plant and bacterial enzyme may have evolved in response to a common selective pressure. This would be consistent with the membrane-bound nature of the plant (Doremus and Jagendorf, 1985) and bacterial enzyme (Karibian, 1978), whereas DHO dehydrogenase appears to be soluble, at least in *D. discoideum* (Jacquet *et al.*, 1985).

Experimental procedures

Strains, media and genetical techniques

Seeds of *Arabidopsis thaliana* (L.). Heynh (Landsberg *erecta* ecotype) originating from J. Giraudat (CNRS, Gif-sur-Yvette) were used. Surface-sterilized seeds were germinated on germination medium containing Murashige and Skoog (1962) salt mixture, 1% sucrose, 100 mg I^{-1} inositol, 1.0 mg I^{-1} thiamine, 0.5 mg I^{-1} pyridoxine, 0.5 mg I^{-1} nicotinic acid and 0.8% agar and were submitted to a 16 h light photoperiod. Seedlings were harvested at the two-leaf stage.

The *E. coli* strain MR32 is a RecA1 derivative of MC1061 (Casadaban and Cohen, 1980) constructed by E. Petrochilo (personal communication). Five *S. cerevisiae* strains were tested for complementation by the *A. thaliana* pool. W303-1B (*mat a, ura3-1, leu2-3,112, his 3-11,15, ade2-101, tryp1-1, can*) was obtained from R. Rothstein, Δ*ura2* is a complete deletion of the *URA2* gene (Bach, personal communication). The three other strains are isogenic to the wild-type strain FL100 (ATCC.28583).

Δ*ura1* was constructed by deletion of the *Hin*dIII–*Bst*EII fragment in the *URA1* gene (Roy, EMBL accession number F32729). *ura4* is a point mutation obtained by UV mutagenesis and *ura5–ura10* is as described in de Montigny *et al.* (1990).

After spheroplast transformation (Hinnen *et al.*, 1978), transformants prototrophic for leucine, histidine, adenine, tryptophan or uracil were selected by plating on omission medium; for example, *HIS*⁺ recombinants were selected on standard minimal medium (Sherman *et al.*, 1986) supplemented with leucine, adenine and tryptophan. The incubation time at 30°C varied from 3 to 10 days. The plasmids were recovered and amplified from individual transformants by subcloning and transformation in the *E. coli* strain MR32 using electroporation. In each selection a control plate without plasmid DNA showed no reversion of the recipient strain, which is in good agreement with the fact that we always found a plasmid in the colonies obtained with transforming DNA.

Sequence analysis

DNA sequences on both strands were determined by the Sanger dideoxynucleotide method using a Sequenase Sequencing kit (USB). We performed either single-strand sequencing from two subclones in M13 or double-strand sequencing on the original p*PyrD ARA* plasmid.

Northern hybridization

Northern blots of poly(A)⁻ RNAs (4 μg) from *A. thaliana* and yeast cells were as described in Sambrook *et al.* (1989) with minor modifications. After denaturation with formaldehyde and electrophoresis in a 1.2% agarose gel, the RNAs were transferred to nitrocellulose and hybridized to ³²P-labelled probes prepared by multiprime labelling. Nonspecific hybridization was washed away by a 30 min treatment at 60°C in 0.1% SDS and 0.2 × SSC.

References

Alderson, A., Sabelli, P.A., Dickinson, J.R., Cole, D., Richardson, M., Kreis, M., Shewry, P.R. and Halford, N.G. (1991) Complementation of *snf1*, a mutation affecting global regulation of carbon metabolism in yeast, by a plant protein kinase cDNA. *Proc. Natl Acad. Sci. USA*, **88**, 8602–8605.

Aruffo, A. and Seed, B. (1987) Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl Acad. Sci. USA*, **84**, 8573–8577.

Becker, D.M., Fikes, J.D. and Guarente, L. (1991) A cDNA encoding a human CCAAT-binding protein cloned by functional complementation in yeast. *Proc. Natl Acad. Sci. USA*, **88**, 1968–1972.

Casadaban, M. and Cohen, S. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli. J. Mol. Biol.* **138**, 179–207.

Doremus, H.D. (1986) Organization of the pathway of *de novo* pyrimidine nucleotide biosynthesis in pea (*Pisum sativum* L. cv Progress no. 9) leaves. *Arch. Biochem. Biophys*. **250**, 112–119.

Doremus, H.D. and Jagendorf, A.T. (1985) Subcellular localization of the pathway of *de novo* pyrimidine nucleotide biosynthesis in pea leaves. *Plant Physiol*. **79**, 856–861.

Frick, M., Neuhard, J. and Kelln, R.A. (1990) Cloning, nucleotide sequence and regulation of the *Salmonella typhimurium pyr* D gene encoding dihydroorotate dehydrogenase. *Eur. J. Biochem*. **194**, 573–578.

Hartl, F.U., Pfanner, N., Nicholson, D.W. and Neupert, W. (1989) Mitochondrial protein import. *Biochim. Biophys. Acta*, **988**, 1–45.

Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Transformation of yeast. *Proc. Natl Acad. Sci. USA*, **75**, 1929–1933.

Jacquet, M., Kalekine, M. and Boy-Marcotte, E. (1985) Sequence analysis of a *Dictyostelium discoideum* gene coding for an active dihydroorotate dehydrogenase in yeast. *Biochimie*, **67**, 583–588.

Karibian, D. (1978) Dihydroorotate dehydrogenase (*E. coli*). In *Methods in Enzymology*, Volume 51 (Hoffee, P.A. and Jones, M.E., eds). New York: Academic Press, pp. 58–63.

Larsen, J.N. and Jensen, K.F. (1985) Nucleotide sequence of the *pyr* D gene of *Escherichia coli* and characterization of the flavoprotein dihydroorotate dehydrogenase. *Eur. J. Biochem*. **151**, 59–65.

Lee, M.G. and Nurse, P. (1987) Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc* 2. *Nature*, **327**, 31–35.

Lew, D.J., Dulic, V. and Reed, S. (1991) Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell*, **66**, 1197–1206.

Lipman, J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. *Science*, **227**, 1435–1441.

Minet, M. and Lacroute, F. (1990) Cloning and sequencing of a human cDNA coding for a multifunctional polypeptide of the purine pathway by complementation of the *ade2–101* mutant in *Saccharomyces cerevisiae. Curr. Genet*. **18**, 287–291.

de Montigny, J., Kern, L., Hubert, J.-C. and Lacroute, F. (1990) Cloning and sequencing of *URA10*, a second gene encoding orotate phosphoribosyl transferase in *Saccharomyces cerevisiae. Curr. Genet*. **17**, 105–111.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant*. **15**, 479–497.

Okayama, H. and Berg, P. (1982) High-efficiency cloning of full-length cDNA. *Mol. Cell. Biol*. **2**, 161–170.

O'Neal, T.D. and Naylor, A.W. (1976) Some regulatory properties of pea leaf carbamoyl phosphate synthetase. *Plant Physiol*. **57**, 23–28.

Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl Acad. Sci. USA*, **85**, 2444–2448.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd Edn. New York: Cold Spring Harbor Laboratory Press.

Schild, D., Brake, A.J., Kiefer, M.C., Young, D. and Barr, P.J. (1990) Cloning of three human multifunctional *de novo* purine biosynthetic genes by functional complementation of yeast mutations. *Proc. Natl Acad. Sci. USA*, **87**, 2916–2920.

Sentenac, H., Bonneaud, N., Minet, M., Lacroute, F., Salmon, J.M., Gaymard, F. and Grignon C. (1992) Cloning and expression in yeast of a plant K^+ transport system. *Science*, **255**, in press.

Sherman, F., Fink, G.R. and Hicks, J.B. (1986) *Laboratory Course Manual for Methods in Yeast Genetics*. New York: Cold Spring Harbor Laboratory Press.

Walther, R., Wald, K., Glund, K. and Tewes, A. (1984) Evidence that a single polypeptide catalyses the two step conversion of orotate to UMP in cells from a tomato suspension culture. *J. Plant. Physiol*. **116**, 301–311.

Wickens, M. and Stephenson, P. (1984) Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent mRNA 3' end formation. *Science*, **226**, 1045–1051.

EMBL Data Library accession number X62909.